CIE independent review report --- BSAI and GOA Flatfish Stock Assessments

Yan Jiao

Prepared for

Center for Independent Experts

Department of Fish and Wildlife Conservation
Virginia Polytechnic Institute & State University
Blacksburg, VA
24061
Phone: 540-2315749
Email: yjiao@vt.edu
http://www.fishwild.vt.edu/faculty/jiao.htm

Contents

Executive Summary

The 2012 stock assessments of four Bering Sea/Aleutian Islands (BSAI) and Gulf of Alaska (GOA) flatfish species were reviewed by a CIE review panel. The four species are GOA northern and southern rock sole, GOA Dover sole, GOA rex sole and BSAI yellowfin sole. The review aims to evaluate the modifications/progress of the stock assessments after 2007, and to ensure that the North Pacific Fishery Management Council (NPFMC) bases its decisions on the best available information when managing these 4 species. The CIE Review Panel met at the Alaska Fisheries Science Center (AFSC), Seattle, Washington during June 11 - 13, 2012. The assessments of the stock done by the stock assessment team were presented to the CIE review panel and the validity of the data, assessment procedures, and results were discussed. The AFSC provided all the necessary logistic support, background information, documents and further data exploration that were requested by the review panel.

The trawl surveys for both the BSAI and GOA were well presented based on the history of the changes in survey frequency, survey spatial coverage, tow duration and sampling of the catches. Some of the concerns from the most recent peer review on the Eastern Bering Sea Crab and Groundfish Bottom Trawl Surveys were discussed again during this review. Concerns discussed included whether the changes in tow duration, the survey depth and distance offshore have been well validated or not.

The Age and Growth Program (AGP) team presented impressive work in which traditional age reading and bomb radiocarbon ^{14}C were combined together with dendrochronology in climate studies. The work from this team provided ageing uncertainty for the four flatfish species to be reviewed, and the growth variation correlated with climate changes.

The effectiveness of the BSAI Amendment 80 (A80) was studied by comparing fleet and productivity seasonable distribution before and after A80. The fishing effort in the final third season increased, the hauls per day decreased after Best Use Cooperative (BUC), the species composition in the productivity by week changed after A80 also. Some analysis on halibut bycatch and pacific cod bycatch studies were presented. These studies indicated that under A80, there are strong incentives for vessels to maximize their expected revenues. Both the multi-species trade-offs in the fishery and seasonal and annual variations of the biological processes of each species can influence the application of A80 and the analysis of the effectiveness of A80.

In general, there was limited reference to model selection uncertainty. Compilations of reports on model selection framework and model selection uncertainty in the future are suggested. Measurements in selecting models can be model goodness-of-fit, model prediction ability, model robustness and fisheries-specific measurements such as retrospective error (Linton and Bence 2011; Jiao et al. 2012). Conducting a maturity study for all the four species is suggested by collecting new data on a regular basis since growth has been observed to vary over time.

The Dover sole assessment was considered preliminary and more model scenarios should be explored in the future before it can be used for providing management reference points. Some key recommendations for Dover sole assessment are summarized below:

- Investigate the combined effect of low sample size and high ageing uncertainty on the stock assessment results.
- Characteristics of the un-sexed individuals in the length-composition need to be explored.
- Year specific age-length curves and length-age transition matrices should be compared with the currently used length-age transition matrix (Stockhausen et al. 2005), to explore the importance of including temporal variation of growth and its influence in stock assessment modeling.
- Estimability of selectivity and natural mortality can be explored in at least 3 ways: model comparison based on goodness-of-fit; simulation study to explore whether selectivity and/or natural mortality of Dover sole is estimable based on its data characteristics (Jiao et al. 2012); and data cloning (Lele et al. 2007; Lele 2010).
- Bayesian estimation does not guarantee global maximum/minimum. Biological meaningful priors for many of the key parameters combine together with the selection of Markov chain Monte Carlo (MCMC) algorithms to help convergence of the model to biologically meaningful estimates.

The rex sole assessment model was an update of the 2009 catch-at-age model and the new model results were consistent with 2009 assessment. The uncertainty of the selectivity estimation prevents the management from using Tier 5 to Tier 3. Some key recommendations for rex sole assessment are summarized below:

- Estimation uncertainty of selectivity needs to be validated through simulation study or data cloning (Lele et al. 2007; Lele 2010).
- Extra exploration of the historical length-composition especially on non-sexed individuals is needed.
- Year specific age-length curves and length-age transition matrices should be compared with the currently used fixed transition matrix to explore the importance of including temporal variation of growth and its influence in stock assessment modeling.

The GOA northern and southern rock sole were assessed together using a sex-specific two species statistical catch-at-age model. The population dynamics of each species were assumed to be independent and parameters of each species are sex-specific. A set of models were explored, which is valuable, but the model exploration was very preliminary and further effort is needed on model development, comparison and selection. Some key recommendations for GOA northern and southern rock sole assessment are summarized below:

- Hierarchical models can be considered in future model development since southern and northern rock soles were considered as one species previously and there are lots of similarities between these two species in biological and fishing processes (Gelman et al. 2004; Jiao et al. 2011).
- Year specific length-at-age is suggested to be used when ageing data are available instead of using a fixed length-at-age curve because the observed length-at-age curves among years are largely different from the currently used sex-specific growth curves (Stark and Somerton 2002).
- The selection of selectivity curves/functions needs to be evaluated through simulation studies and based on a clear model comparison/ selection framework.
- Maximizing posterior likelihood (MPLE) was used to estimate parameters and to compare the seven alternative models. It is useful to provide a comparison of the results when MPLE and MCMC are used in solving the same model with the same parameterization and prior assumptions.
- Simulation studies can be used to explore the robustness of the methods (both survey based and model based approaches should be considered) on survey relative abundance index standardization under situations when gear, trawl duration and trawling spatial coverage changed (Yu 2010).
- Spatial variation of the fishery / survey over time should be explored to validate the effectiveness of the survey design and estimator of abundance.
- A simulation study on how small sample size of age-composition influences the stock assessment uncertainty is suggested. The exploration should provide a scientific basis for the suggestion of future biological sampling.

The assessment of BSAI yellowfin sole was considered to represent the best scientific information available for the stock assessment of this species although improvements or adjustments in model structure development are possible. This BSAI yellowfin sole was assessed based on a statistical catch-at-age model with fixed M, temperature dependent q, using directly observed weight-at-age from each year's biological sampling, and time varying selectivity. The recommended base model was chosen by the assessment team after assessing a set of statistical catch-at-age models with multiple submodels of weight-at-age, natural mortality and catchability. Selectivity is assumed to follow a logistic curve but changes every year. Some key recommendations for BSAI yellowfin sole assessment are summarized below:

- Compare performance of constant q and temperature dependent q with other formulas of temperature dependent q, such as quadratic forms.
- A simulation study is suggested to explore whether selectivity needs to be modelled as time varying. If time varying is found to be necessary, factors that cause the selectivity to vary largely needs to be explored.
- Using observed weight-at-age data directly is suggested but for age groups in some years with small sample size, smoothing among years or model based approach may be used to avoid measurement uncertainty caused by small sample size or samples from limited spatial locations.

- Further research is suggested on how environmental variables influence life history traits such as growth, maturity, recruitment/productivity, and fishing processes such as selectivity and catchability.
- Preliminary analysis on the relationship between catch-per-unit of effort (CPUE) and environmental factors, such as water temperature, current direction/strength and habitat types etc., should be conducted to see whether the relationship in the Northern Bering Sea is different from the relationship in the current survey area based on the existing one year survey.
- A further computer intensive simulation study can also be conducted to evaluate whether survey with and without the Northern Bering Sea area would result in the same abundance index after standardization, i.e., whether spatial incompleteness in survey coverage further influence the development of the indicator of the whole population (Yu 2010).

1. BACKGROUND

This report reviews the 2012 stock assessments of four flatfish species distributed in the Bering Sea/Aleutian Islands (BSAI) and Gulf of Alaska (GOA) at the request of the Center for Independent Experts (CIE). I was provided with draft stock assessment reports and web access to relevant files and documents (Appendix 1) and participated in the Stock Assessment Review Meeting. Extra documents were provided during the review upon request from the CIE peer review panel (Appendix 1).

The 4 species that underwent review are GOA northern and southern rock sole, GOA Dover sole, GOA rex sole and BSAI yellowfin sole. These 4 species are important fisheries resources in Alaska and are key components of the BSAI and GOA ecosystem (North Pacific Fishery Management Council 2011). According to the CIE project description, "The flatfish stock assessments routinely undergo thorough review by the AFSC, the North Pacific Fisheries Management Council's Groundfish Plan Teams and Scientific and Statistical Committee, and members of the public." The last CIE peer review of these stock assessments was in 2007 and there have been several modifications and developments in their assessment methodologies since then. Considering the importance of these species in the Alaska ecosystem and the important role of stock assessments in their management, the CIE review in 2012 is timely and beneficial.

Prior to the presentations on the stock assessments of the four species under review, an overview of the fisheries in BSAI and GOA and their management processes and management tiers were introduced by Dr. Wilderbuer. Presentations on the bottom survey trawl conducted in GOA, AI and BS and the observer programs were also provided (Appendix 1). These surveys and observer programs provided crucial data used in the stock assessments of Alaska flatfishes.

For the four species under review, the current stock assessment model used, the biological references points estimated, and the recommended Tier for management were presented by the stock assessment team (see Agenda in Appendix 2). Discussions on the quality of the data, the appropriateness of the model equations and error structures, key parameters, and the estimation algorithms were made throughout the review.

The quality of the data for each species was discussed and comments and suggestions are listed under ToRs. Positive evidences include 1) cohort signals of many of the flatfish species can be seen from the bar plot and bubble plot of the survey catch-at-age/length (Wilderbuer 2012; Nichol 2012); 2) spatial patterns of the fishery are similar to the survey except little fishery effort occurs in the Eastern GOA.

Because the time limitation of conducting four species stock assessment review in three days, discussions on model development and model selection were not thorough. During the review meeting, the stock assessment team was always available when required for further discussion, additional data exploration and clarification, and clarification of how the Alaska flatfish Terms of Reference (ToR) were addressed. A conclusion was then drawn on which management Tier to use as the basis for management of the specific fishery. The ToRs for each stock were quickly reviewed in the afternoon of June 13 to ensure they had been fully addressed. Recommendations from NPFMC SSC reports and minutes of Plan Teams of GOA and BSAI groundfish were both reviewed to determine the extent to which they had been addressed.

As a CIE reviewer, my duty was to evaluate the stock assessments of the four species with respect to the ToRs, which are attached in Annex 2 of the Appendix 2. This report provided the findings of the independent review that is undertaken by me in accordance with the CIE Statement of Work.

2. ROLE OF INDIVIDUAL REVIEWER IN THE REVIEW ACTIVITIES

My role as a CIE independent reviewer was to conduct an impartial and independent peer review in accordance with the SoW and the predefined ToRs herein.

About two weeks before the meeting, assessment documents and supporting materials (NPFMC, 2011a and 2011b) were made available to the review panel via email and a http website by Dr. Wilderbuer. I read all the documents that I received prior to the review.

The CIE peer review meeting was held at the AFSC, Seattle, Washington, from June 11-13, 2012. The meeting followed the "tentative agenda (Appendix 2)" of the CIE review. The meeting was open and was organized constructively. On the morning of June 11 before the meeting, the assessment review committee met with Dr. Wilderbuer and the stock assessment team to discuss the meeting agenda, reporting requirements, and meeting logistics. During the meeting, all documents were made available electronically through Dropbox and ftp website.

Presentations were given during the review according to the agenda to provide the CIE reviewers the background information on the data used in the stock assessment models, the model development, and the recommended management tiers. I was actively involved in the discussion during the presentations by 1) listening to the presentations carefully, making notes on the points that are not included or not clearly stated in the documents provided prior to the meeting; 2) asking questions for clarification on the data usage and model development; 3) making comments and providing possible alternative solutions to questions arising during the meeting.

After the peer review meeting, I summarized the findings and recommendations according to the predefined ToRs. This review report is formatted according to my interpretation of the required format and content described in Annex 1 of Appendix 2.

3. SUMMARY OF FINDINGS AND RECOMMENDATIONS FOR ToRs

Below I provide the summary of findings of each ToR in which the weaknesses and strengths are described and conclusions and recommendations in accordance with the ToRs.

3.1 GOA Dover sole TOR

CIE Reviewers shall evaluate the current model assumptions and make recommendations for improvements thereof, including:

3.1.1. Use of age data, including:

a. use of age composition data

The percentage disagreement of the age reading is as high as 75% (Helser 2012). The influence of ageing uncertainty on the age-length transition matrices is not clear. Also the influence of the small sample size is not clear, especially with the combination of ageing uncertainty.

For years with ageing data, the year specific age-length curves and age-length transition matrix should be compared with the currently used transition matrix (Stockhausen et al. 2005). This comparison will help to explore the necessity of using temporal variation of growth and its influence on stock assessment modeling, such as the influence on the estimates of selectivity, the biomass/ abundance and the biological/management reference points. A thorough evaluation of the impact of ageing uncertainty and sample size on the stock assessment results is suggested for future stock assessment.

b. appropriateness of age range and binning

The current model combined ages 35-39 into age bin of age 35 and used age 40 as a +group. The observed age composition (Fig 5.9b in NPFMC, 2011a) indicated that both age 35 bin and age 40+ can be substantial. Extra simulation or at least sensitivity analysis should help to explore the appropriateness of the current age bin and +group. Since multinomial distribution was used for age composition data and observations of zeros were mainly in 2 years (1984 and 1987), I suggest that age bin of age 35-39 be separated as individual bins. At very least a sensitivity analysis should be explored.

c. estimation of size-at-age relationship and variability (external vs. internal to model)

External is more appropriate to avoid further complexity. The estimability/ identifiability of selectivity is a big concern for the stock assessment of this species. The variation of length-at-age may be driven by many factors but the direct observation should give us the exact temporal length-at-age (including its uncertainty) and the year-specific length-age transition matrix. Special attention is needed, however, when direct observation is from a small sample size. A standard can be established on how to deal with this situation, such as using model derived length-at-age or using multi-year smoothed values when sample size is less than a certain number.

d. inclusion of ageing error

The inclusion of ageing error needs to be investigated given the fact that the percentage disagreement of the age reading is as high as 75% (Helser 2012). A simulation should be done to evaluate the combined effect of ageing error with small sample size in the age composition on the stock assessment.

3.1.2. Use of size data, including:

a. use of survey size composition data

The survey size composition data are appropriately used in the current stock assessment. Extra sensitivity analysis on the use of different length bins would help diagnose the robustness of the stock assessment results in the length bin assumptions.

b. use of fishery size composition data

Quite a high percentage (higher than 50%) of fishes in the length composition sampling were not sexed and it is not clear how the unsexed individuals influence the sex-specific length composition. Characteristics of the un-sexed individuals in the length-composition need to be investigated, such as the size frequency of the un-sexed individuals. With such investigation, we can judge whether the fishery size composition data should be used or not, or we can investigate the influence of the unsexed individuals in size composition on the stock assessment results.

3.1.3. The number and functional forms of estimated selectivity curves, including:

The current stock assessment that we reviewed did not incorporate alternative functional forms of selectivity curves. Previous explorations of this question were included in Stockhausen et al. (2009). The findings below are based on both the onsite discussion and these in Stockhausen et al. (2009).

a. fitting different selectivity functions to data from different survey years based on survey depth coverage

Estimability of selectivity may be explored in at least 3 ways: 1) model comparison among model scenarios of non-sex specific selectivity-at-length and sex-specific selectivity; 2) simulation study to explore whether selectivity of Dover sole is estimable based on the data characteristics of Dover sole (Jiao et al. 2012); and 3) data cloning (Lele et al. 2007; Lele 2010).

b. types of selectivity curves considered

The current model used logistic curves to model selectivity of both sexes and surveys with different coverage. Age-specific selectivity curves can be used to explore the selectivity patterns at age. For example, age-specific selectivity can be modeled as a random walk process and then the selectivity pattern estimated can further be described using an equation (Linton and Bence 2011).

c. use of age-based vs. size-based selectivity curves

Both fish behavior and model comparison need to be explored before one either age-based or size-based selectivity curves are used in the future. Fish behavior may be explored to determine whether size-specific or sex-specific selectivity is more appropriate through field monitoring and/or expert knowledge.

d. allowing for annual variability in fishery selectivity

Both the review panel and the stock assessment team agreed that allowing annual variability in fishery selectivity is an appropriate model alternative to be explored in the future.

e. use of size-based selectivity curves for survey data based on trawl net catchability experiments

Somerton et al. (2007) was provided during the peer review meeting. The paper conducted net efficiency experiment to explore catchability and selectivity over size of 4 flatfish species. Based on the importance of the survey catchability in this stock assessment and in the 4 species that we reviewed, I support that this is worthwhile to be investigated in the future. Results should be applied carefully however. Both fishing behavior and gear performance need to be explored to interpret the results. Also, results from these experiments may be used as informative priors on the selectivity curves in the statistical catch-at-age model later, instead of using the results to fix the selectivity curves.

3.1.4. Fixing (and updating) the natural mortality rate based on Hoenig (1983)

Natural mortality, selectivity and catchability were all important parameters in the statistical catch-at-age models. The stock assessment team explored alternative functional forms of selectivity in the 2009 stock assessment (Stockhausen et al. 2009), and will explore natural mortality rates in the future. Simulation studies are suggested to be used to explore questions including 1) whether natural mortality should be fixed, 2) whether natural mortality can be estimated with the life history based empirical natural mortality estimates (such as Hoenig 1983; Pauly 1980) used as priors, and 3) explore time varying natural mortality and the factors that cause natural mortality to vary over time and space (Aanes et al. 2007; Jiao et al. 2012).

3.1.5. Model convergence diagnostics

The stock assessments results of 2011 and 2009 were largely different, thus whether the Markov chain Monte Carlo (MCMC) converged in the 2011 model is of high concern to the stock assessment team. Because of time limitation, none of the MCMC chains was requested or explored during the review meeting. However, based on the discussion, commonly used methods (Cowles and Carlin 1996) such as trace plots, Gelman-Rubin statistics, and autocorrelation plot have been used by the stock assessment team. Bayesian parameter estimation does not guarantee global maximum /minimum, and in many cases we look for biologically meaningful estimates. Biological meaningful priors for many of the key parameters, such as M, selectivity-at-age, and population size from the survey combine together with the selection of MCMC algorithms to help convergence of the model to biologically meaningful estimates.

3.2 GOA rex sole TOR

CIE Reviewers shall evaluate, and make recommendations for improvements on, the current approach to determining stock status and future harvest reference points (ABC and OFL).

The rex sole assessment model is an update of the 2009 catch-at-age model and the model results are consistent with 2009 assessment also. Given the uncertainty in both the estimated selectivity pattern and the lack of reasonable quality recruitment dynamics model, I agree with the recommendation of the stock assessment team on the management Tier, i.e., Tier 5.

The uncertainty of the selectivity estimation prevents management from using Tier 5 to Tier 3. Estimation uncertainty of selectivity needs to be validated through modeling, for example, simulation study or data cloning (Lele et al. 2007; Lele 2010). Both fisheries and biological processes should be discussed to further explore the selectivity pattern. The simulation may be used to diagnose 1) whether biomass can be estimated with reasonable precision given the uncertainty of the selectivity estimation; 2) whether the selectivity estimation is of reasonable precision, so that it can be used in deriving $F_{40\%}$ and $F_{35\%}$ or F_{msy} and B_{msy}.

Other findings and suggestions on the model specification and data usage:

a. The specification of selectivity using normalized fishery selectivity is confusing. The plot of the selectivity indicated that the posterior runs of selectivity were rescaled, so that the maximum of the posterior mean selectivity-at-age is 1. Since MCMC was used in this study, for each run, selectivity was modeled as a logistic function with max (selectivity-at-age) =1, then the joint posterior runs of selectivity shouldn't be rescaled. The posterior mean selectivity-at-age should be actually lower than 1.

b. Catch composition of bycatch and catch from directed fishery may be added in the document later, to improve understanding of both the data and the management implications.

c. It is not clear why some of the fishes sampled in the length-composition are not sexed and how it may influence the sex-specific length composition and further influence the stock assessment. Extra exploration on the historical length-composition is needed on these un-sexed individuals. It may help to diagnose whether it is one of the reasons for the difficulty in estimating stable selectivity results.

d. Age-length samples from the surveys should be used to develop length-age conversion matrices of their corresponding years and then compare these with the currently used constant conversion matrix.

e. Age-length data can be used to not only explore the growth variation of the rex sole over time but also help explore how using a constant age-length conversion matrix may influence the stock assessment results, such as cohort signal, abundance/biomass estimation, and biological reference points estimation.

3.3 GOA northern and southern rock sole TOR

3.3.1 Evaluation, findings and recommendations of the analytical approach (application of a statistical AD Model Builder (ADMB) integrated catch-age model) used to assess stock status and estimation/presentation of uncertainty.

A base model was presented, which is a sex-specific two species statistical catch-at-age model. The population dynamics of each species are assumed to be independent and parameters of each species are sex-specific. There were discussions on whether these two species should be separated considering the complexity/simplicity of the model. The stock assessment team plotted the length composition of the 2 species over time upon request and found that there are clear differences between the two species. A further genetic analysis based on tissue samples from multiple locations would help address this question also.

The observed length-at-age curves among years are largely different from the currently used sex-specific growth curves (Stark and Somerton 2002). We recommend that year specific length-at-age be used when ageing data are available instead of using a fixed length-at-age curve. Also factors that may influence the growth variation need to be explored in the future, so that projections of the model can incorporate the growth variation in the short future.

The stock assessment document and the presentation in the meeting only presented the uncertainty of the estimated population size/spawning stock size. One concern that arose from both the review panel and the stock assessment team is the sex-specific selectivity-at-length/age of the two species. Whether species specific and sex-specific selectivity-at-length should be simplified to non-sex specific, for example, may be explored through simulation study.

It was recognized that a tremendous amount of effort has been used to explore the model uncertainty by comparing seven models (with different assumptions of natural mortality and selectivity). These efforts were considered quite valuable. However both the stock assessment team and the review panel realized that the exploration on the model uncertainty is preliminary. More effort is needed in the future to address the selection of the model and provide the appropriate models for the stock assessment of the two species.

Maximizing posterior likelihood was used to estimate parameters (MPLE) and in comparing the seven alternative models. It is useful to provide a comparison of the results when MPLE and MCMC are used in solving the same model with the same parameterization and prior assumptions (Magnusson et al. 2012). When both process error and observation or measurement error are considered, parameter estimation is a problem and Bayesian approach has been recommended as an effective method (Millar and Meyer 2000; Calder et al. 2003). A recent study also suggests that MCMC is a better choice in solving statistical catch-at-age models (Magnusson et al. 2012). Also a clear model comparison framework needs to be described in the future when multiple models are considered and compared.

3.3.2 Evaluation findings and recommendations on quality of input data and methods used to process them for inclusion in the assessment (specifically fishery and survey data).

The observed length-at-age curves among years are largely different from the currently used sex-specific growth curves (Stark and Somerton 2002). It is recommended that year specific length-at-age be used when ageing data are available instead of using a fixed length-at-age curve.

Although the overall ageing uncertainty is low for both northern and southern rock sole compared to other flatfish species (Helser 2012), exploration on how the ageing uncertainty influences key parameters of population dynamics and the parameters of management importance would be beneficial in the future.

Concerns arose about the survey abundance estimation, and changes in the catchability and selectivity when gears, trawl duration and trawling spatial coverage changed. Simulation studies can be used to explore the robustness of the methods on the survey relative abundance index standardization under the situations of the changes listed above (Yu 2010). Both survey based and model based approaches can be used to estimate survey abundance and a simulation study can be used to explore the appropriateness of these approaches for the specific species based on the historical survey data.

Concerns also arose about whether the survey abundance data captured the cohort signals. Data exploration such as bubble plots and bar plots of the survey abundance-at-age should be provided to better answer this question.

Data of survey biomass and survey/fishery length composition are not species specific before the 1996 survey. Also the trawling period changed from 1996 and onward. An alternative model may be used that does not incorporate data before 1996. The influence of not using the pre-1996 data on the parameters and abundance estimates can be investigated and may shed light on future model/data usage when more survey years are accumulated.

Change of the length-at-age relationship over time has been observed and the review team suggested that the year-specific length-at-age be used in the future. The maturity-at-age is fixed in the model based on Stark and Somerton (2002). Based on the fact that the observed length-at-age relationship over time is changing, maturity-at-age is very possible to vary over time also. Future biological sampling and model construction should consider time varying or year-specific maturity-at-age/length.

Spatial variation of the fishery / survey over time should be explored in the future to help understand the spatial distribution patterns of the species. This type of exploration should influence/validate the effectiveness of the survey design and estimator of abundance estimation.

Sample size of the age composition sampling tends to be low in many years. A simulation study on how sample size of age-composition influences the stock assessment uncertainty should be evaluated in the future. The evaluation should provide a scientific basis for the suggestion of future biological sampling.

3.3.3 Recommendations for further assessment improvements for management in both the long and short term.

Hierarchical models should be considered in future model development. Since southern and northern rock soles were considered as one species previously and there are many similarities between these 2 species in their biological processes, using hierarchical modelling allows the stock assessment model to borrow strength from the data sets of both populations, instead of assuming that both populations are independent (He and Sun 2000; Jiao et al. 2011). Multi-level priors also result in robust parameter estimates (Roberts and Rosenthal 2001; Gelman et al. 2004; Jiao et al. 2011).

Model selection uncertainty and model comparison need to be further explored, and a well-designed model selection framework is needed. Measurements in selecting models can be model goodness-of-fit, model prediction ability, model robustness, and fisheries-specific measurements such as retrospective error (Jiao et al. 2012; Linton and Bence 2011).

3.4 BSAI yellowfin sole TOR

3.4.1 Evaluation of the analytical approach (application of a statistical ADMB integrated catch-age model) and model assumptions used to assess stock status and stock productivity.

This stock was assessed using a set of statistical catch-at-age models and the population dynamics were calibrated using estimated survey biomass and age compositions both from fishery and surveys. The submodels explored include four types of weight-at-age models, two natural mortality (M) models (one is to use fixed M=0.12), two catchability (q) models (one is the temperature dependent catchability). Maturity is assumed to be fixed based on a study from Nichol (1995). Selectivity is assumed to follow a logistic curve but changes every year. The recommended base model from the stock assessment team incorporate a fixed M, temperature dependent q, using directly observed weight-at-age from each year's biological sampling, and time varying selectivity. The stock assessment report indicated that the Maximum Likelihood Estimation approach was used to estimate parameters and compare models; however, many of the results were shown as posterior distributions from MCMC approach. Stock recruitment (SR) relationship was analyzed out of the statistical catch-at-age model analysis and SR relationships were explored based on 3 time blocks.

These investigations on different submodel scenarios were considered very valuable, although the presentation of them in both the stock assessment model and in the presentation were not in enough details and were not fully discussed because of the time limitation during the review meeting. A model selection framework and model selection uncertainty evaluation is suggested for future stock assessment.

The parameters estimated by the model are presented below:

Fishing mortality	Selectivity	Survey catchability	Year class strength	Spawner - recruit	Total
60	240	2	58	2	362

Concerns arose on the model complexity such as using temperature dependent catchability and time varying selectivity. I support the research effort to explore changes of catchability and selectivity over time and space if possible. Studies on other species found that q correlated with water temperature (Chifamba 2000). Here only one extra parameter was added to the model using temperature dependent q. I also would like to suggest that the stock assessment team explore some other formulas of temperature dependent q, such as quadratic forms. Quantification of how extra model complexity increases the model goodness-of-fit is suggested to be added in the future stock assessment. Selectivity was found to be largely different among years. I understand that this is the model recommended based on model goodness-of-fit; however, factors that cause the large variation of selectivity-at-age among years need to be explored for the benefit of 1) understand the fish/fishing behavior; and 2) better estimation of both BRPs and future stock projection. I support the selection of using observed weight-at-age data directly, but for age groups in some years with small sample size, smoothing or a model based approach should be used to avoid measurement uncertainty caused by small sample size or samples from limited spatial locations.

Yellowfin sole's somatic growth and catchability were found to be correlated with both water temperature and sea surface temperature. Maturity-at-age may be influence by water temperature or vary among years also. More updated biological sampling on maturity would benefit the stock assessment. I also support further effort on the exploration on how environmental variables influence life history traits such as growth, maturity, recruitment/productivity, and fishing processes such as selectivity and catchability.

3.4.2 Evaluation of the implications of using the Northern Bering Sea research results as an index of abundance if yellowfin sole increasingly occupy this area with changing climate.

Analysis on the relationship between catch-per-unit of effort (CPUE) and environmental factors, such as water temperature, current direction/strength and habitat types etc., should be conducted to see whether the relationship in the northern Bering Sea is different from the relationship in the current survey area based on the existing one year survey. Also age-composition and length-composition in both the Northern Bering Sea and the Southern Bering Sea can be compared to evaluate the age/size groups that occupy this area. In general, I support the idea of continued survey in the Northern Bering Sea and the development an index of abundance for future stock assessment. A further computer intensive simulation study should also be conducted to evaluate whether survey with and without the Northern Bering Sea area would result in the same abundance index after standardization, i.e., whether spatial incompleteness in survey coverage (only part of the spatial area that the population distributed is surveyed) further influence the development of the indicator of the whole population (Yu 2010).

3.4.3 Determination of whether the assessment represents the best available science for the stock assessment of BSAI yellowfin sole, including considerations of fishery rationalization on timing and selectivity of fishery.

A separate presentation (Haynie 2012) was provided on the BSAI fishery rationalization study but there are no specific documents provided on BSAI yellowfin sole for this stock assessment review. Both the yellowfin sole catch cumulative distribution function plot and the bar plot of production by week showed that the seasonal pattern of yellowfin sole catch in BSAI changed after A80. The night fishing intensity during weeks 13-20 when yellowfin sole is one of the major target species decreased to a degree (reduction of night fishing of 9.1% in 2008 and 4.7% for 2009 and 2010). The decrease of the night fishing intensity largely decreases the bycatch of halibut because the ratio of halibut/yellowfin sole per hour of fishing is 30-83% higher at night time.

My comments on selectivity functions used in the stock assessment can be seen from ToR 3.4.1. I support the research effort to explore changes of selectivity over time. Also, how different selectivity functions improve model goodness-of-fit needs to be quantified in the stock assessment document, so that we can understand the rationale of using year-specific selectivity curves, which occupy 66% of the total number of parameters.

I consider the assessment represents the best scientific information available for the stock assessment of BSAI yellowfin sole although improvements or adjustments in model structure development are possible. The review panel considered the BSAI yellowfin sole assessments sufficient to provide the basis for the management of this fishery.

4. SUGGESTIONS FOR IMPROVEMENTS OF NMFS REVIEW PROCESS

The current review process is very well designed. It can be further improved if the powerpoints used in the review meeting can be distributed to the review panel a few days earlier before the meeting, and if a follow-up review can be conducted in the near future. Full Bayesian analysis is time consuming and suggestions on model structural variation do not seem appropriate to request and hard to finish in one to two nights. The stock assessment review and discussion should be implemented more effectively by this extra follow-up review.

5. Acknowledgements

I would like to thank all the Stock Assessment Team members contributing to the meeting for their informative presentations on the stock assessments of Alaska flatfish and for providing helpful responses to the review panel's questions. Many thanks also to Bering Sea survey and the GOA survey groups, and the Age and Growth Program and other scientists at the meeting for their presentations and their contribution to the discussions throughout the meeting. Special thanks also go to other members of the review panel, Drs. Sven Kupschus and Kevin Stokes for productive discussions on the assessments.

6. References

Aanes, S., Engen, S., Saether, B-E., and Aanes, R. 2007. Estimation of the parameters of fish stock dynamics for catch-at-age data and indices of abundance: can natural and fishing mortality be separated? Canadian Journal of Fisheries and Aquatic Sciences, 64: 1130–1142.

Calder, C., Lavine, M., Muller, P., and Clark, J.S. 2003. Incorporating multiple sources of stochasticity into dynamic population models. Ecology. 84: 1395-1402.

Chifamba, P. 2000. The relationship of temperature and hydrological factors to catch per unit effort, condition and size of the freshqater sardine, *Limnothrissa miodon* (Boulenger), in Lake Kariba. Fisheries Research. 45: 271-281.
Cowles, M.K., and Carlin, B.P. 1996. Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review, Journal of the American Statistical Association. 91: 883-904.
Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. 2004. Bayesian Data Analysis. Chapman & Hall. 668pp.
He, Z., and D. Sun. 2000. Hierarchical Bayes estimation of hunting success rates. Biometrics 5:223–236.
Jiao, Y. Cortes, E., Andrews, K. and Guo, F. 2011. Poor-data and data-poor species stock assessment using a Bayesian hierarchical approach. Ecological Applications 21:2691-2708.
Jiao, Y., Smith, E., O'Reilly, R., and Orth, D. 2012. Modeling nonstationary natural mortality in catch-at-age models: an example using the Atlantic weakfish (Cynoscion regalis) fishery. ICES Journal of Marine Science 69:105-118.
Lele, S. R., Dennis. B., and Lutscher, F. 2007. Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods . Ecology Letters. 10: 551-563.
Lele, S.R. 2010. Model complexity and information in the data: Could it be a house built on sand? Ecology. 91: 3493-3496.
Linton, B.C., and Bence, J.R. 2011. Catch-at-age assessment in the face of time-varying selectivity. ICES Journal of Marine Science, 68: 611–625.
Magnusson, A., Punt, A.E., Hilborn, R. 2012. Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and MCMC. Fish and Fisheries. DOI: 10.1111/j.1467-2979.2012.00473.x
Millar, R.B., and Meyer, R. 2000. Bayesian state–space modeling of age-structured data: fitting a model is just the beginning. Canadian Journal of Fisheries and Aquatic Sciences, 57: 43–50.
Roberts, G.O. and Rosenthal, J.S. 2001. Infinite hierarchies and prior distributions. Bernoulli. 7:453-471.
Stockhausen, W.T., M.E. Wilkins and M.H. Martin. 2009. 5. Gulf of Alaska Deepwater Flatfish. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. pp. 495-563. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510.
Yu, H. 2010. Spatial and temporal population dynamics of yellow perch (Perca flavescens) in Lake Erie. PhD. Thesis, Virginia Tech.

Appendix 1: Bibliography of materials provided for review

Draft Stock Assessments:

North Pacific Fishery Management Council (NPFMC). 2011a. Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea/Aleutian Islands Region.

Introduction and overview of stock assessment.

Chapter 4: Assessment of the Yellowfin sole stock in the Bering Sea and Aleutian Islands

North Pacific Fishery Management Council (NPFMC). 2011b. Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska.

Introduction and overview of stock assessment

Chapter 4: Assessment of the Shallow-water Flatfish Complex in the Gulf of Alaska for 2012

Chapter 5: Assessment of the Deepwater Flatfish Stock in the Gulf of Alaska

Chapter 6: Assessment of the Rex Sole Stock in the Gulf of Alaska

Presented Materials:

A'mar, Teresa. 2012. CIE review of the GOA northern and southern rock sole stock assessment. Alaska Fisheries Science Center.

Haynie, Alan. 2012. The rationalization of the BSAI amendment 80 Fleet. Economics and Social Sciences Research Program. Alaska Fisheries Science Center.

Helser, Thomas. 2012. Flatfish age determination at AFSC (including research). Alaska Fisheries Science Center.

Nichol, Dan and others.2012. Eastern Bering Sea shelf bottom trawl survey of groundfish and invertebrate resources. Groundfish Assessment Program RACE Division. Alaska Fisheries Science Center.

Palsson, Wayne. 2012. Gulf of Alaska and Aleutian Island Bottom Trawl surveys. Alaska Fisheries Science Center.

Stockhausen, William. 2012. Dover Sole. Alaska Fisheries Science Center.

Stockhausen, William. 2012. GOA Rex Sole. Alaska Fisheries Science Center.

Thompson, Lisa. 2012. Fisheries monitoring and analysis division, North Pacific Groundfish Observer Program.

Wilderbuer, Tom and Ianelli, Jim. 2012. CIE review of the Bering Sea/Aleutian Islands yellowfin sole stock assessment.

Wilderbuer, Tom. 2012. Overview of flatfish fisheries and management. Alaska Fisheries Science Center.

Other Materials upon request from the CIE panel:

A'mar, Teresa. 2012. Survey length data for GOA northern and southern rock sole.

CIE review of the bottom trawl surveys. 2012. Reports from Yong Chen, Jon Helge Vølstad, and N.G. Hall.

North Pacific Fishery Management Council. 2011. DRAFT REPORT of the SCIENTIFIC AND STATISTICAL COMMITTEE to the NORTH PACIFIC FISHERY MANAGEMENT COUNCIL.

North Pacific Fishery Management Council. 2011. Minutes of the Bering Sea Aleutian Islands Groundfish Plan Team.

North Pacific Fishery Management Council. 2011. Minutes of the Gulf of Alaska Groundfish Plan Team.

North Pacific Fishery Management Council. 2011. Minutes of the Joint Plan Teams for the Groundfish Fisheries of the Gulf of Alaska (GOA) and Bering Sea Aleutian Islands (BSAI).

Somerton, D.A., Munro, P.T., and Weinberg, K.L. 2007. Whole-gear efficiency of a benthic trawl for flatfish. Fish. Bull. 105:278–291.

Spencer, Paul and Ianelli, James. 2012. Temperature-dependent survey catchability for eastern Bering Sea flatfish. (manuscript in preparation)

Stark, J.W. and Somerton, D.A. 2002. Maturation, spawning and growth of rock soles off Kodiak Island in the Gulf of Alaska. Journal of Fish Biology. 61, 417–431.

Stockhausen, William. 2012. Dover sole age compositions bubble plots for both males and females. Alaska Fisheries Science Center.

Appendix 2: Statement of Work for Dr. Yan Jiao

Attachment A: Statement of Work for Dr. Yan Jiao

External Independent Peer Review by the Center for Independent Experts

Peer Review of the BSAI and GOA flatfish stock assessments

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description: The Alaska Fisheries Science Center (AFSC) requests a Center of Independent Experts (CIE) review of 4 Bering Sea/Aleutian Islands (BSAI) and Gulf of Alaska (GOA) flatfish stock assessments. They include: GOA northern and southern rock sole, GOA Dover sole, GOA rex sole and BSAI yellowfin sole. The BSAI and GOA flatfish resources are large, subject to significant fisheries and are key components of the BSAI and GOA ecosystems. The flatfish stock assessments routinely undergo thorough review by the AFSC, the North Pacific Fisheries Management Council's Groundfish Plan Teams and Scientific and Statistical Committee, and members of the public. However, the BSAI and GOA flatfish stock assessments have not had the benefit of a CIE review since 2007. Since 2007, several modifications to existing assessment and projection models have been implemented, and a new assessment for Gulf of Alaska northern and southern rock sole has been developed. These innovations have not been reviewed by the CIE. The Alaska Fisheries Science Center desires an independent peer review of these stocks to assess the quality of the assessments and to ensure that the North Pacific Fishery Management Council bases its decisions on the best available information. Therefore, a CIE review in 2012 would be timely. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge and recent experience in the application of stock assessment, including population dynamics, survey methodology, estimation of parameters in complex nonlinear models, and the AD Model assessment program in particular. Reviewers should also have experience conducting stock assessments for fisheries management. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Seattle, Washington tentatively during June 11-13, 2012.

Statement of Tasks: Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/sponsor.html).

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Other Tasks – Contribution to Summary Report: Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate in the panel review meeting in Seattle, Washington during June 11-13, 2012.
3) In Seattle, Washington during June 11-13, 2012 as specified herein, conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than July 9, 2012, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and CIE Regional Coordinator, via email to David Die ddie@rsmas.miami.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

May 1, 2012	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
June 1, 2012	NMFS Project Contact sends the CIE Reviewers the pre-review documents
June 11-13, 2012	Each reviewer participates and conducts an independent peer review during the panel review meeting
July 9, 2012	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
July 23, 2012	CIE submits CIE independent peer review reports to the COTR
July 30, 2012	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on changes. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) each CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) each CIE report shall address each ToR as specified in **Annex 2**,

(3) the CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Program Manager, COTR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-427-8155

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
shivlanim@bellsouth.net Phone: 305-383-4229

Roger W. Peretti, Executive Vice President
Northern Taiga Ventures, Inc. (NTVI)
22375 Broderick Drive, Suite 215, Sterling, VA 20166
RPerretti@ntvifederal.com Phone: 571-223-7717

Key Personnel:

Tom Wilderbuer, NMFS Project Contact
NMFS Alaska Fisheries Science Center
7600 Sand Point Way NE., Seattle, WA 98115-6349
tom.wilderbuer@noaa.gov Phone: 206-526-4224

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Tentative Terms of Reference for the Peer Review

2012 CIE Review for selected Gulf of Alaska and Bering Sea flatfish

GOA Dover sole TOR

CIE Reviewers shall evaluate the current model assumptions and make recommendations for improvements thereof, including:

1. Use of age data, including:
    a. use of age composition data
    b. appropriateness of age range and binning
    c. estimation of size-at-age relationship and variability (external vs. internal to model)
    d. inclusion of ageing error
2. Use of size data, including:
    a. use of survey size composition data
    b. use of fishery size composition data
3. The number and functional forms of estimated selectivity curves, including:
    a. fitting different selectivity functions to data from different survey years based on survey depth coverage
    b. types of selectivity curves considered
    c. use of age-based vs. size-based selectivity curves
    d. allowing for annual variability in fishery selectivity
    e. use of size-based selectivity curves for survey data based on trawl net catchability experiments
4. Fixing (and updating) the natural mortality rate based on Hoenig, 1983.
5. Model convergence diagnostics

GOA rex sole TOR

CIE Reviewers shall evaluate, and make recommendations for improvements on, the current approach to determining stock status and future harvest reference points (ABC and OFL).

GOA northern and southern rock sole TOR

1. Evaluation, findings and recommendations of the analytical approach (application of a statistical ADMB integrated catch-age model) used to assess stock status and estimation/presentation of uncertainty.

2. Evaluation findings and recommendations on quality of input data and methods used to process them for inclusion in the assessment (specifically fishery and survey data).
3. Recommendations for further assessment improvements for management in both the long and short term.

BSAI yellowfin sole TOR

Evaluation of the analytical approach (application of a statistical ADMB integrated catch-age model) and model assumptions used to assess stock status and stock productivity.
Evaluation of the implications of using the Northern Bering Sea research results as an index of abundance if yellowfin sole increasingly occupy this area with changing climate.
Determination of whether the assessment represents the best available science for the stock assessment of BSAI yellowfin sole, including considerations of fishery rationalization on timing and selectivity of fishery.

Annex 3: Tentative Agenda

CIE Flatfish assessment review

NMFS Alaska Fisheries Science Center
7600 Sand Point Way NE, Building 4
Seattle, Washington

AGENDA June 11-13, 2012

Monday June 11th

9:00	Welcome and Introductions, adopt agenda	**Sandra**
9:15	Overview (species, biology, surveys, fishery, catch levels, ABCs, TACs, bycatch)	**Tom**
10:00	Bering Sea trawl survey	**RACE Division**
10:30	Gulf of Alaska trawl survey	**RACE Division**
11:00	Coffee break	
11:20	Observer Program	**FMA Division**
11:50	Age Determination	**Delsa and Beth**
12:30	Lunch	
1:30	Effect of rationalization on flatfish fisheries	**REFM Economic subtask**
2:30	GOA rex sole	**Buck**

Tuesday June 12th

9:00	Gulf of Alaska Dover sole	**Buck**
11:00	Coffee break	
11:20	Gulf of Alaska Dover sole (continued)	**Buck**
12:30	Lunch	
1:30	Gulf of Alaska northern and southern rock sole	**Teresa**

Wednesday June 13th

9:00	Bering Sea yellowfin sole	**Tom and Jim**
11:00	Coffee break	
11:20	Bering Sea yellowfin sole (continued)	**Tom and Jim**
12:30	Lunch	
1:30	CIE panel discussion (assessment authors will be available)	

Appendix 3:

Panel membership or other pertinent information from the panel review meeting

CIE Reviewers

Sven Kupschus
Yan Jiao
Kevin Stokes

Presenters from Stock Assessment Team

Teresa A'mar
James Ianelli
Sandra Lowe
Buck Stockhausen
Tom Wilderbuer

Presenters from Alaska Survey, Observer, and Age and Growth Programs

Dan Nichol	Bering Sea survey program
Wayne Palsson	Gulf of Alaska survey program
Tom Helser	Age and growth program
Lisa Thompson	Observer program
Anne Hollowed	